

Jeremy B. Sporn
Federal Defenders of Eastern Washington and Idaho
306 East Chestnut Avenue
Yakima, Washington 98901
(509) 248-8920

Attorneys for the Defendant

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF WASHINGTON
The Honorable SALVADOR J. MENDOZA, JR.

United States of America,

Plaintiff,

v.

James D. Cloud,

Defendant.

Case No. 1:19-cr-2032-SMJ-1

**Motion to Exclude or Limit
Toolmark Identification Evidence**

Yakima - With Oral Argument

July 20, 2021, at 10:00

I. Introduction

We are witnessing a sea change in how courts approach and rule on matters relating to forensic science. Nowhere is this more apparent than in the field of firearms and toolmark identification, where an examiner purports to “match” fired ammunition components to an individual firearm. The last several years have featured a paradigm shift in how trial courts approach firearms identification evidence, from almost unquestioned deference towards outright skepticism. On the heels of multiple reports from the scientific community that cast doubt on the foundational validity or reliability of the practice, and pursuant to their heightened gatekeeping function under *Daubert*, courts have sharply restricted the testimony of firearms examiners as to a claimed identification or match.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 Here, the government has charged James Cloud with participating in the murders of
2 five people, and associated other serious crimes, including carjacking and kidnapping. If
3 proven, he will likely spend the remainder of his life in prison. Short of a death penalty
4 prosecution, the charges are among the most serious this Court is likely to see. But the
5 government seeks to prove its case against Mr. Cloud in part through the type of firearms
6 and toolmark identification evidence that lacks sufficient indicia of reliability to convict
7 someone of *any* criminal offense, let alone charges of such gravity. And there is a substantial
8 possibility that, unless excluded or limited, the jury will accord undue weight to such
9 evidence presented through the sheen of cocksure, expert testimony.

11 Consistent with this Court's gatekeeping function under Federal Rule of Evidence
12 702 to first ensure the reliability of evidence the jury will consider to determine his fate, Mr.
13 Cloud moves to exclude or, in the alternative, limit firearms identification evidence and
14 testimony. At a minimum, the Court should not permit the examiner to opine that
15 recovered shell casings were fired from a particular firearm. For the following reasons, the
16 motion should be granted.

18 **II. Factual Background and Proposed Reports and Testimony**

19 The Court is already familiar with the basic outline of the government's charges and
20 its version of events. Through extensive briefing provided as part of Mr. Cloud's motion to
21 suppress or exclude eyewitness identification, and the testimony provided over two days at
22 that hearing, as well as other motion practice to date, the parties have provided the Court
23 with an extensive recitation of the government's investigation, which there is little need to
24 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
25

repeat here. This motion concerns physical evidence recovered from several different locations indicated in the chart below. First, shell or cartridge casings collected from JC's Medicine Valley residence, and two vehicles (both Honda Civics) found in two other locations¹. It also concerns a .22 caliber Ruger rifle recovered from an irrigation canal close to where Morris Jackson and Natasha Jackson were arrested, and which Mr. Jackson purportedly discarded in the canal shortly before apprehension.

<u>Item Number</u>	<u>Description</u>	<u>Source</u>
1B131, 1B134, 1B136, 1B146	.22 LR cartridge casings	Medicine Valley residence
1B198, 1B200, 1B205(A), 1B205(B), 1B205(C), 1B210	.22 LR cartridge casings	1989 blue Honda Civic
005 (1B282), 006 (1B283), 007	.22 LR cartridge casings	2004 blue Honda Civic
64 (1B11)	Ruger .22 rifle	Irrigation canal

¹ One Honda Civic, the blue 1989 model, was found in a remote area, off Highway 97 on the Yakama reservation, about 10-12 miles southwest of Toppenish on June 8, 2019. The government suspects that James Cloud was associated with that vehicle. The other 2004 Honda Civic was associated with two concert goers from California who went missing, and whose remains were recovered in August 2020 in a remote area off Highway 97, well southeast of the Medicine Valley residence. On June 8, 2019, the blue 2004 Honda Civic was found in an orchard adjacent to and just east of 8100 Lateral B Road in Wapato, also on the Yakama reservation. The government advised at the last status conference that on the basis of information in its possession as of that date (May 18, 2021), it did not intend to charge James or Donovan Cloud in connection with the disappearance of the two Californians.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 Kristen Drury, a forensic lab supervisor with the Yakima Police Department,
2 examined the 17 shell casings itemized above. Following her comparison analysis, wherein
3 the 17 casings were “visually examined and microscopically compared to each other,” Ms.
4 Drury opined that twelve cartridge cases [**1B131, 1B134, 1B136, 1B146, 1B198, 1B200,**
5 **1B205(A-C), 1B210, and Items 005 and 006]** were all “identified as having been fired in the
6 same .22 Long/Long Rifle caliber firearm.” Exhibit A. Drury also found that the five
7 cartridge cases comprising **Item 007**, were “eliminated as having been fired in same .22
8 Long/Long Rifle . . . due to differences in class characteristics.” *Id.* However, she found
9 that those five shell casings had all been fired from the same .22 Long Rifle as one another.
10

11 The casings, as well as additional items were analyzed subsequently by FBI forensic
12 examiner Michael Van Arsdale. In February 2020, Mr. Van Arsdale reported that he “would
13 prefer to directly compare the test-fired casing with the recovered casings rather than rely on
14 NIBIN,” and the FBI sent Mr. Van Arsdale the recovered casings. Following his analysis,
15 Van Arsdale opined that the same twelve “cartridge cases² were identified as having been
16 fired in the Item 64 rifle [the .22 Ruger recovered from the irrigation canal]”. Exhibit B.
17

18 The government’s Rule 16 expert notice indicates, *inter alia*, that Van Arsdale “will
19 testify that bullets were fired from the Ruger rifle and examined. Arsdale examined several
20

21
22
23 ² As to the bullets or bullet fragments recovered from the bodies of the five victims, Mr. Van
24 Arsdale opined that they either lacked sufficient corresponding microscopic marks of value to permit a
25 conclusion as to whether they were fired from the .22 Ruger Rifle or from the same firearm, that they “bear
no marks of value suitable for comparison purposes,” or that they were otherwise unsuitable for
comparison purposes. Exhibit B.

1 bullets (Items 109-119) that were provided to him. Arsdale conducted a comparison and
 2 determined that the bullets (Items 109-119) had been fired from the Ruger rifle.”
 3 Elsewhere, the government similarly described Van Arsdale’s testimony, noting that based
 4 on his education, training, experience and examinations in the instant case, he “will offer the
 5 following opinions: Items 108-119 [described *supra*] (cartridge cases) had been fired from
 6 the Ruger [rifle referenced earlier in the summary].” By this motion, Mr. Cloud seeks to
 7 exclude such testimony, whether proffered by Drury³, Van Arsdale, or both.
 8

9 **III. Burdens and Evidentiary Gatekeeping**

10 At the outset, it is the government’s burden to show the reliability of its proposed
 11 expert testimony. That is, as the proponent of its proposed toolmark identifications, the
 12 *government* bears the burden of demonstrating its admissibility by a preponderance of the
 13 evidence. *See, e.g., Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593 n.10 (1993); *Cayuga*
 14 *Indian Nation v. Pataki*, 83 F. Supp. 2d 318, 322 (N.D.N.Y. 2000) (“In the context of the
 15 admissibility of expert opinion evidence under *Daubert*, . . . it is the proponent’s burden . . .
 16 to establish admissibility, rather than the opponent’s burden to establish inadmissibility.”)
 17 (quoting and citing *Graham v. Playtex Products, Inc.*, 993 F. Supp. 127, 129 (N.D.N.Y. 1998)
 18 and *Lust v. Merrell Dow Pharmaceuticals*, 89 F.3d 594, 598 (9th Cir. 1996)); *see also United States v.*
 19 *Wells*, 2014 U.S. Dist. LEXUS 200824, at *5 (D. Alaska. Feb. 5, 2014) (the “proponent bears
 20
 21

22
 23
 24 ³ Drury’s opinion differs in that she purports to believe that the relevant shell casings match each
 25 other, i.e., that they were fired from the same weapon, and not they were fired from the particular .22 Ruger
 rifle that was recovered from the irrigation canal and test-fired, as does Van Arsdale’s.
 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 the burden of showing that the expert's findings are based on sound science [, which]
2 requires some objective, independent validation of the expert's methodology”). Mr. Cloud
3 does not have the burden of disproving or raising sufficient doubt as to firearms
4 identification, or demonstrating that it is not reliable.

5
6 Federal Rules of Evidence 702 provides that expert testimony is admissible when
7 qualified “by knowledge, skill, experience, training, or education,” but only where:

- 8 (a) the expert's scientific, technical, or other specialized knowledge will help the trier of
9 fact to understand the evidence or to determine a fact in issue;
10 (b) the testimony is based on sufficient facts or data;
11 (c) the testimony is the product of reliable principles and methods; and
12 (d) the expert has reliably applied the principles and methods to the facts of the case.

13 Rule 702 reflects the codification of the Supreme Court’s decision in *Daubert*, and
14 charges the district court, as the evidentiary gatekeeper, “with the task of ensuring ‘that any
15 and all scientific testimony or evidence admitted is not only relevant, but reliable.’” *United*
16 *States v. Freeman*, 498 F.3d 893, 901 (9th Cir. 2007) (quoting *Daubert*, 509 U.S. at 589). This
17 requires trial courts to make a “preliminary assessment of whether the reasoning or
18 methodology underlying the testimony is . . . valid and of whether that reasoning or
19 methodology properly can be applied to the facts in issue.” *Freeman*, 498 F.3d at 901 (quoting
20 *Daubert*, 509 F.3d at 592-93).

22 District courts are to exercise this gatekeeping function for all forms of expert
23 testimony, not just on scientific matters, since the linchpin of *Daubert* and Rule 702 is
24 reliability. See *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 147-49 (1999); *United States v.*
25 *Motion to Exclude or Alternatively Limit Toolmark Identification Evidence*

1 *Hermanek*, 289 F.3d 1076, 1093 (9th Cir. 2002) (requiring district courts to assure that expert
2 testimony “rests on a reliable foundation and is relevant to the task at hand”) (quoting
3 *Daubert*, 509 U.S. at 292-93)); *see also Freeman*, 498 F.3d at 904 (“District courts have the
4 continuing responsibility of acting as the vigilant gatekeepers of expert testimony to ensure
5 that it is reliable”). Science may not have the monopoly on reliability, as far as expert
6 testimony goes, but judges must ensure that testimony as to technical or other specialized
7 knowledge rests on a similarly reliable foundation and is applied reliably.

8
9 In performing this function and deciding whether to admit expert testimony, courts
10 are guided by several non-exhaustive *Daubert* factors, including: (1) whether the theory or
11 technique can be, and has been, tested; (2) whether the theory or technique has been
12 subjected to peer review and publication; (3) the known or potential error rate of the
13 technique; (4) whether standards exist and are maintained to control the technique’s
14 operation; and (5) the degree to which the theory or technique is generally accepted in the
15 relevant scientific community. *Daubert*, 509 U.S. at 593-94; *see also United States v. Prime*, 431
16 F.3d 1147, 1152 (9th Cir. 2005). As is the case with any type of evidence, even if expert
17 testimony is both relevant and reliable, it is nevertheless subject to exclusion under Rule 403
18 where its probative value is substantially outweighed by the danger of unfair prejudice. *See*
19 *United States v. Vallejo*, 237 F.3d 1008, 1016-1021 (9th Cir. 2001).

20
21 The district court’s gatekeeping thus involves far more than applying a rubber
22 stamp on esoteric knowledge. To wit, “nothing in either *Daubert* or the Federal Rules of
23 Evidence requires a district court to admit opinion evidence that is connected to existing
24 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
25

1 data only by the *ipse dixit* of the expert. A court may conclude there is simply too great an
 2 analytical gap between the data and the opinion proffered.” *General Electric Co. v. Joiner*, 536
 3 U.S. 136, 146 (1997). Put differently, the trial court’s task “requires more than simply taking
 4 the expert’s word for it.” *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1319 (9th Cir.
 5 1995). An “expert’s bald assurance of validity is not enough. Rather, the party presenting
 6 the expert must show that the expert’s findings are based on sound science, and this will
 7 require some objective, independent validation of the expert’s methodology.” *Id.* at 1316.
 8 As Judge Gertner noted in ruling on a similar challenge: “*Daubert* plainly raised the standard
 9 for existing, established fields, inviting a reexamination even of generally accepted venerable,
 10 technical fields. Refusing to do so would be equivalent to grandfathering old irrationality.”
 11 *United States v. Green*, 405 F. Supp. 2d 104, 118 (D. Mass. 2005).
 12

13 **B. A Hearing, Rather than Cross-Examination at Trial, is Necessary**

14 The preliminary assessment referenced above frequently follows what has become
 15 known as a *Daubert* hearing, an opportunity to hear from the potential expert, as well as the
 16 challenging party’s proposed expert, outside the presence of the jury, and well before the
 17 jury is empaneled. A defense opportunity to *voir dire* the government witness at trial would
 18 be, at least in this case, an inadequate mechanism for the Court to make its preliminary
 19 assessment as to the reliability of firearms identification evidence. Given the complexity of
 20 the subject matter, recent developments in toolmark litigation (which typically have followed
 21 the sort of hearing that Mr. Cloud requests), and the high stakes for both Mr. Cloud and the
 22
 23
 24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 government, the Court should explore these issues more fully through testimony at an
2 evidentiary hearing.

3 This case and the proposed expert testimony are wholly unlike the sorts of cases
4 where courts have dispensed with the formalities of a *Daubert* hearing. *See, e.g., United States*
5 *v. Alatorre*, 222 F.3d 1098, 1105 (9th Cir. 2000) (affirming district court’s denial of *Daubert*
6 hearing and admission of officer’s’ expert testimony as to street value of marijuana and
7 distribution quantities); *United States v. Brown*, 2017 U.S. Dist. LEXIS 126504, at *9-12 (C.D.
8 Cal. Aug. 7, 2017) (denying *Daubert* hearing and permitting experienced DEA Agent to
9 testify as to street prices of cocaine and that Los Angeles was a hub for distribution). Those
10 examples are not difficult for a jury to understand, and do not pose significant challenges for
11 courts in making a preliminary finding of reliability. But the complexity and nuance of
12 firearms identification is far different from that type of testimony – particularly in the age of
13 *CSI*, where juries might be inclined to accept forensic practices at face value and without
14 much scrutiny
15

16
17 Similarly, cross-examining the expert at trial as to particular issues surrounding
18 foundational validity, or presenting a defense expert to the jury would not be sufficient, in
19 this case. For one, many of the issues raised herein go to admissibility, not weight. Second,
20 by the time the defense might have such an opportunity, it might be too late; the persuasive
21 effects of “scientific” testimony might be too firmly implanted in the minds of the jurors.
22 As another judge, faced with a similar motion and request, wisely put it:
23

24 While cross-examination may often play such a role, this discipline and the
25 disputes surrounding it seem far too complex for a series of questions on cross-
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 examination to allow a full understanding of the limitations of the field. Indeed, a
2 full exploration of the issues surrounding the reliability of this evidence in the
3 present case required several days of testimony from multiple expert witnesses,
4 close evaluation of numerous applied-science studies, exploration into the studies'
5 design and methodology and the problems arising therefrom, and advocacy by
6 counsel on each side specially tasked with litigating forensic science issues. It
7 would be fanciful to conclude that the normal adversarial process would enable a
8 lay jury to adequately understand these issues, and it is similarly unrealistic to
9 conclude that the average attorney in the average trial would be able to raise these
10 issues in front of the jury in this fashion, particularly when this issue would be one
11 among many issues to be presented to the jury in a trial.

12 *United States v. Tibbs*, 2019 D.C. Super. Ct. LEXIS 9, at *79-80 (Sept. 5, 2019).

13 Denying a hearing and requiring Mr. Cloud to address these issues through cross-
14 examination at trial would be, in effect, to “abandon the gatekeeping function altogether, for
15 Rule 702 clearly contemplates *some* degree of regulation of the subjects and theories about
16 which an expert may testify.” *United States v. Ruvalcaba-Garcia*, 923 F.3d 1183, 1189 (9th Cir.
17 2019) (quoting *Kumho Tire*, 526 U.S. at 158-59 (Scalia, J., concurring) and *Estate of Barabin v.*
18 *AstenJohnson, Inc.*, 740 F.3d 457, 464 (9th Cir. 2014)). The Ninth Circuit has faulted trial
19 courts for failing to assess the validity or methodology of proposed expert testimony, and for
20 “delegating” that role to the jury. *See Ruvalcaba-Garcia*, 923 F.3d at 1189 (citing cases).
21 Consistent with the court’s duty to make an explicit finding of reliability before admitting
22 such testimony, we respectfully request a hearing. *See id.* at 1189-90 (discussing “abdication”
23 of courts’ gatekeeping function).

24
25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

IV. Toolmarks and AFTE Theory of Identification

At its core, the premise of toolmark analysis or firearms identification is straightforward: microscopic imperfections or marks⁴ acquired during the manufacture and use of a firearm, are thereafter transferred to bullets and shell casings as they come into contact with interior components and are projected out the barrel of the gun during the firing process. These markings, the theory goes, are sufficiently unique so as to permit ascertainable identification, or source attribution, to that individual firearm. *See United States v. Shipp*, 422 F. Supp. 2d 762, 769-70 (E.D.N.Y. 2019) (quoting *United States v. Otero*, 849 F. Supp. 2d 425, 428-28 (D. N.J. 2012)) (providing a helpful description of toolmark identification); President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (September 20, 2016) [the "PCAST Report," attached as Exhibit C]. Yet the simplicity of the theory does not confer reliability in the abstract or validity in application. In other words, that the very idea is logical, and that we can easily enough wrap our heads around it, does not mean there is a scientific or reliable basis to it. Indeed, the validity of the doctrine depends on two unfounded premises: (1) that an individual firearm will leave distinct and unique markings on a bullet fragment or shell casing, and that (2) trained examiners are capable of recognizing and differentiating among such impressions so as to make source attributions.

⁴ These markings or surface contours are generally known as striations [striae] or impressions, and also as grooves, lands and by other terminology, although these terms do not all mean the same thing.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 Most firearms examiners belong to the Association of Firearm and Tool Marks
2 Examiners [“AFTE”], which is essentially a professional association or trade organization.
3 As relevant here, it is *not* a scientific entity. Toolmark analysis is based on what the AFTE
4 describes as its subjective, non-scientific “theory of identification”:

- 5 1. The theory of identification as it pertains to the comparison of toolmarks enables
6 opinions of common origin to be made when the unique surface contours of two
7 toolmarks are in ‘sufficient agreement’.
- 8 2. This ‘sufficient agreement’ is related to the significant duplication of random
9 toolmarks as evidence by the correspondence of a pattern or combination of patterns
10 of surface contours. Significance is determined by the comparative examination of
11 two or more sets of surface contour patterns comprised of individual peaks, ridges and
12 furrows. Specifically, the relative height or depth, width, curvature and spatial
13 relationship of the individual peaks, ridges and furrows within one set of surface
14 contours are defined and compared to the corresponding features in the second set of
15 surface contours. Agreement is significant when the agreement in individual
16 characteristics exceeds the best agreement demonstrated between toolmarks known to
17 have been produced by different tools and is consistent with agreement demonstrated
18 by toolmarks known to have been produced by the same tool. The statement that
19 “sufficient agreement” exists between two toolmarks means that the agreement of
20 individual characteristics is of a quantity and quality that the likelihood another tool
21 could have made the mark is so remote as to be considered a practical impossibility.
- 22 3. Currently the interpretation of individualization/identification is subjective in nature,
23 founded on scientific principles [this is, of course, very much in dispute] and based on
24 the examiner’s training and experience⁵.

25 The differences between class characteristics, sub-class characteristics, and what are
known as individual characteristics are critical to understanding key issues in toolmark
analysis and firearms identification. AFTE defines “class characteristics” as “measurable

⁵ AFTE Theory of Identification, The Association of Firearm and Tool Mark Examiners,
available at <https://afte.org/about-us/what-is-afte/afte-theory-of-identification> (last visited May 24, 2021).
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 features of a specimen which indicate a restricted group source. They result from design
 2 factors and are determined prior to manufacture⁶.” *See also Shipp*, 422 F. Supp. 2d at 770
 3 (describing class characteristics as “gross features common to most if not all bullets and
 4 cartridge cases fired from a type of firearm”). These include the caliber of the firearm, for
 5 example, or the rifling pattern used in a class of firearm, including the direction of the
 6 “twist” and the number of lands and grooves in the barrel. Class characteristics “generally
 7 originate in the product design phase intentionally and are considered common to every
 8 member of a relatively large group of items or product [and] encompass intentional design
 9 characteristics.” *Declaration of William A. Tobin* [attached as Exhibit D], at 13.

11 Subclass characteristics, meanwhile, are “features that may be produced during
 12 manufacture that are consistent among items fabricated by the same tool in the same
 13 approximate state of wear. These features are not determined prior to manufacture and are
 14 more restrictive than class characteristics⁷.” *See also Shipp*, 422 F. Supp. 2d at 770 (explaining
 15 that sub-class characteristics are “produced incidental to manufacture but apply only to a
 16 subset of the firearms produced, for example, as may occur when a batch of barrels is
 17 formed by the same irregular tool”). As to individual characteristics, AFTE defines them as

21 ⁶ https://afte.org/uploads/documents/AFTE_Glossary_Version_6.110619_DRAFT_PDF.

22 ⁷ https://afte.org/uploads/documents/AFTE_Glossary_Version_6.110619_DRAFT_PDF.
 23 Subclass characteristics are “within the particular class of product (*e.g.*, .38 caliber revolver), including
 24 striations and impressions that derive incidentally from production processes. Subclass characteristics are
 25 not present on the product/component by design, but rather are fortuitously acquired during the
 manufacturing process from a forming/shaping tool that can leave virtually identical markings on large
 numbers of the products produced, including firearms components [.]” Exhibit D, at 14.
 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

“marks produced by the random imperfections or irregularities of tool surfaces. These random imperfections or irregularities are produced incidental to manufacture and/or caused by use, corrosion, or damage. They are unique to that tool to the practical exclusion of all other tools.” *See also United States v. Adams*, 444 F. Supp. 2d 1248, 1252-53 (D. Ore. 2020) (providing overview of the steps and processes of standard AFTE methodology); *United States v. Tibbs*, 2019 U.S. Dist. LEXIS 9, at *4-8 (D.C. Super. Ct. Sept. 5, 2019). This motion and the salient controversies surrounding toolmark analysis, mostly concern the uniqueness and discernibility of individual characteristics, for that is what permits an examiner to claim a “match” or identification to a particular firearm.

V. Recent Toolmark Developments in Science, Litigation and Research

A. Scientific Reports Cast Doubt

In the last dozen years, highly respected academics and scientists have weighed in on toolmark analysis several times, and their findings have significantly altered the landscape when it comes to admitting firearms identification testimony in criminal proceedings. Mainstream scientists or the “relevant scientific community” have made generalized findings as to the validity of firearms identification in two reports in particular. First, in August 2009, the National Research Council, a division of the National Academy of Sciences, published a report entitled, *Strengthening Forensic Science in the United States: A Path Forward*⁸. NRC criticism

⁸ This “NRC Report” is quite lengthy and is available at <https://www.ojp.gov/pdffiles1/nij/grants/228091.pdf>. The 2009 report actually drew on another NRC report from the year before, which had reached the conclusion that “the validity of the fundamental Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 centered around the fact that AFTE guidelines as to firearms identification were highly
2 subjective, and did not feature a “precisely defined process” or “provide a specific protocol”.
3 NRC Report, at 155. It noted that AFTE guidance “does not even consider, let alone
4 address, questions regarding variability, reliability, repeatability, or the number of
5 correlations needed to achieve a given degree of confidence.” *Id.* Reviewing the relevant
6 theory and evidence, the NRC report found that an examiner’s identification “remains a
7 subjective decision based on unarticulated standards and no statistical foundation for the
8 estimation of error rates.” *Id.* at 153-54. In other words, what firearms examiners claim to
9 do is inconsistent with the scientific method, and the report recommended additional studies
10 to “make the process of individualization more precise and repeatable”. *Id.* at 154. As it
11 stood, however, “sufficient studies have not been done to understand the reliability and
12 reproducibility of the methods” involved. *Id.*

15 Then, in 2016, the President's Council of Advisors on Science and Technology
16 weighed in, in even more critical fashion in a lengthy report entitled, *Forensic Science in*
17 *Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*. The report deals with
18 7 “feature comparison” methods, including firearms identification, and the scientific criteria
19 to establish foundational validity⁹, which include reproducible procedures for identifying,
20

21
22 assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully
23 demonstrated.” National Research Council, *Ballistic Imaging*, at 81 (2008).

24 ⁹ The PCAST Report [Exhibit C, at 43] defined “foundational validity” to mean “the scientific
25 standard corresponding to the legal standard of evidence being based on 'reliable principles and methods'”.
As relevant here, the PCAST Report [Exhibit C, at 47], likewise discusses three essential criteria or indicia of
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 comparing and matching [source attribution] sample features, as well as considering
 2 empirically-based error rates and sensitivity. The latter is particularly important for
 3 subjective methods like toolmark analysis, whose foundational validity required empirical
 4 testing. PCAST is critical of examiners' reliance on training and experience to somehow
 5 validate, as a scientific matter, the practice of firearms identification: "Nothing—not
 6 training, personal experience nor professional practices—can substitute for adequate
 7 empirical demonstration of accuracy." Exhibit C, at 46, 60-61. The report criticized the
 8 circular reasoning behind AFTE's theory of identification, and its reliance on the vaguely-
 9 defined "sufficient agreement" standard: "The 'theory' states that an examiner may
 10 conclude that two items have a common origin if their marks are in 'sufficient agreement,'
 11 where 'sufficient agreement' is defined as the examiner being convinced that the items are
 12 extremely unlikely to have a different origin¹⁰." Exhibit C, at 104.

13
 14
 15 The PCAST report further reviewed a number of studies, detailing the inherent
 16 flaws in design and methodology and their diminished, minimal value towards
 17

18
 19
 20 reliability – **repeatability** (when examiners obtain the same results when dealing with the same samples),
 21 **reproducibility** (when *different* examiners obtain the same results from the same samples), and **accuracy**
 (the degree to which the examiner obtains correct results as to ground truth).

22 ¹⁰ The circularity inherent in this premise is best understood by removing superfluous words,
 23 such that an "examiner may conclude that two items have a common origin if [...] the examiner is
 24 convinced that the items are extremely unlikely to have a different origin." Whatever else that is – an echo
 chamber in the examiner's head or as a "fallacy of presumption" – it is not scientific or objective. *See*
Adams, 444 F. Supp. 2d at 1262 (finding that the "AFTE 'sufficient agreement' standard is a tautology that
 doesn't *mean* anything") (emphasis in original).

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

demonstrating accuracy, finding that such studies may “substantially underestimate the false positive rate” and did not replicate casework. Exhibit C, at 106-08. The key take-away was as follows:

Although firearms analysis has been used for many decades, only relatively recently has its validity been subjected to meaningful empirical testing. Over the past 15 years, the field has undertaken a number of studies that have sought to estimate the accuracy of examiners’ conclusions. While the results demonstrate that examiners can under some circumstances identify the source of fired ammunition, many of the studies were not appropriate for assessing scientific validity and estimating the reliability because they employed artificial designs that differ in important ways from the problems faced in casework.

Exhibit C, at 106. “Because there has been only a single appropriately designed study¹¹, the current evidence falls short of the scientific criteria for foundational validity.” Exhibit C, at 111. The PCAST report indicates that “[t]he scientific criteria for foundational validity require more than one such study, to demonstrate reproducibility.” *Id.* at 150. The report thus discusses the need for more appropriately-designed studies, and development of less subjective standards. *Id.* at 113-14.

The PCAST Report does not debunk toolmark identification entirely, or prohibit the possibility that an individual examiner may correctly attribute bullet fragments or shell

¹¹ This was a “black-box” study under the auspices of the Ames National Laboratory, *see* David P. Baldwin et al., *A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons* (2014) and attached as Exhibit E. PCAST defined a “black-box study” as an “empirical study that assesses a subjective method by having examiners analyze samples and render opinions about the origin or similarity of samples.” Exhibit C, at 48-49. It is so named because the “method must be evaluated as if it were a “black box” in the examiner’s head. *Id.* at 5-6. In firearms identification studies, the ground truth, as to whether a comparison sample is a match, is known by the experimenter, and such studies typically feature an “open-set” design for making independent decisions, rather than a more simple “closed-set” or paired matching exercise.

casings to a particular firearm. Rather, it calls into question whether, from a scientific point of view, the discipline has been shown reliable enough to achieve foundational validity. But the distinction is key, in light of which party has the burden of demonstrating reliability. If firearms examiners and AFTE proponents cloak themselves in the mantle of science for purposes of demonstrating the reliability of the field or “imbue[ing] the practice with an undeserved perception of infallibility and reliability” [Exhibit D, at 17], then they must be evaluated by scientific criteria. Yet the implicit critique in the PCAST Report is that you cannot use science when it is convenient to do so, and then forsake its strictures when they are unachievable.

The scientific community has spoken and not in favor of the reliability of firearms identification. Just the opposite. We recognize that the PCAST report is not Gospel. However, the Department of Justice’s critical response¹², issued in the waning days of the Trump administration (after the insurrection at the Capitol on January 6), seemingly rejects scientific principles and focuses more on semantic, ideologically-driven distinctions between comparisons and measurements, the exclusivity of certain scientific validation testing, and the fallibility or imprecision of error rates¹³. Ironically, it ends up making the point that

¹² Available at <https://www.justice.gov/olp/page/file/1352496/download>

¹³ For a detailed and thoughtful criticism of DOJ’s response, see Thomas D. Albright, *The U.S. Department of Justice Stumbles on Visual Perception*, Proceedings of the National Academy of Sciences [attached as Exhibit F]; *Letter by Christina Swarns to Department of Justice*, dated February 4, 2021 [attached as Exhibit G], at 6 (concluding that DOJ’s statement “uses misleading rhetorical pretzels rather than scientific evidence to refute a scientific report; and all for the purpose of preserving specious litigation strategies”); Maneka Sinha, *The Trump DOJ Snuck in One Last Effort to Push Junk Science in Court*, SLATE, Feb. 4, 2021, available at Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 toolmark identification is extra-scientific, and not capable of the kinds of testing that is a
 2 predicate for admissibility. Not surprisingly, the strongest detractors of the PCAST Report
 3 were prosecutors and the AFTE, as well as the Organization of Scientific Area Committees,
 4 Firearms and Toolmark Subcommittee. In contrast, the Council that authored the PCAST
 5 Report was composed primarily of scientists, academics and business persons, described as
 6 the “Nation’s leading scientists and engineers.” Exhibit C, at iv-ix. The defense bar largely
 7 stayed out of the way, and PCAST’s legal advisors were primarily federal judges and
 8 academics. *Id.* at viii-ix. They had far less skin in the game than the PCAST critics.

10 **B. Judges Have Begun Listening to and Relying on Science**

11 The prosecution’s typical response to similar defense motions is that Courts have
 12 always admitted this type of evidence. And there is some truth in that: “For most of the
 13 twentieth century, courts generally allowed firearm examiners to testify, without many
 14 restrictions, that a bullet found at the scene of a crime was fired from a particular gun.”
 15 *United States v. Romero-Lobato*, 379 F. Supp. 3d 1111, 1116 (D. Nev. 2019) (citation omitted);
 16 *United States v. Glynn*, 578 F. Supp. 2d 567, 569 (S.D.N.Y. 2008) (“[F]or many decades
 17 ballistics testimony was accepted almost without question in most federal courts in the
 18 United States”). But “that’s the way we’ve always done it” is not a persuasive justification
 19
 20

21
 22
 23 <https://slate.com/news-and-politics/2021/02/trump-doj-forensic-science-pcast.html>. In spite of its title,
 24 Sinha’s piece, much like Swarns’ letter and Albright’s position paper, is largely apolitical and non-ideological,
 25 and engages on the lack of scientific rigor underlying DOJ’s position, and its ideological, outcome-driven
 underpinnings.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 for doing something, let alone something reliable enough to justify its admission in a
2 criminal trial, and especially when new evidence comes along that calls into question whether
3 the discipline has any foundational validity.

4 Justice Holmes famously said that “[m]ost of the things we do, we do for no better
5 reason than that our fathers have done them or our neighbors do them [. . . and called it]
6 revolting to have no better reason for a rule of law than that so it was laid down in the time
7 of Henry IV. It is still more revolting if the grounds upon which it was laid down have
8 vanished long since, and the rule simply persists from blind imitation of the past.” Oliver
9 Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457, 468-69 (1897). That would
10 appear to be the case in this arena, but things have progressed, and are continuing to change.
11 And evidence or rules derived from the time of Henry IV might be *less* problematic because
12 it was the age of superstition. Here, instead, we have pseudo-science masquerading as
13 reliable evidence. The veneer, or superficial markers of science and reliability can sometimes
14 be more pernicious than outright superstition or sheer supposition¹⁴.

17 However, in spite of the inauspicious history when it comes to forensic toolmark
18 examinations, in recent years, and especially in the aftermath of the PCAST Report, both
19 federal and state court decisions have reached something of a tipping point. The trend is
20

21
22 ¹⁴ For instance, in recent years, creationists have sought to teach creationism through so-called
23 “intelligent design” and to buttress creationist beliefs with the trappings of scientific thought, in order to
24 dispute evolutionary theory. But wrapping up the same product in the language of science does not change
the underlying or core unscientific premise, or somehow make it more reliable. And seeking a scientific
imprimatur but without the rigor of scientific methodology is a meaningless exercise.

definitely in favor of limitation or restriction, if not yet outright exclusion. These cases include, but are not limited to the following representative samples:

United States v. Green, 405 F. Supp. 2d 104, 124 (D. Mass. 2005) (limiting expert's testimony to observations of similarities and differences between toolmarks, and excluding opinion that shell casing was or was likely to have been fired from particular firearm);

United States v. Adams, 444 F. Supp. 3d 1248, 1267 (D. Ore. 2020) (limiting testimony to description of similarities as to caliber, directional twist and numbers of lands/grooves, and excluding opinion that casings found at scene of shooting matched or were fired from Taurus handgun recovered from defendant's closet¹⁵);

People v. Ross, 68 Misc. 3d 899, 918 (Supreme Court, Bronx County, June 30, 2020) (limiting testimony to evidence of class characteristics and the inability to exclude firearm in question as source of shell casings);

United States v. Davis, Case No. 18-cr-11, 2019 WL 4306971, at *7-8 (W.D. Va. Sept. 11, 2019) (permitting examiners to testify about perceived similar and consistent markings but not that such marks indicate a match or that the casings were fired by same gun, that individual firearms have a "signature" toolmark, or to opine as to any level of confidence in identification)

United States v. Tibbs, 2019 D.C. Super. LEXIS 9, at * 80-81 (Sept. 5, 2019) (limiting examiner's testimony to inability to exclude firearm in question as source of casing found at shooting, and noting that "any statements by the expert involving more certainty regarding the relationship between a casing and a firearm would stray into territory not presently supported by reliable principles and methodology");

United States v. Medley, Case No. 17-cr-242-PWG, Transcript of Motions Hearing, at 54 (D. Md. Apr. 24, 2019) (precluding expert's opinion that different cartridges were fired from same gun or expression of any level of confidence);

United States v. Shipp, 422 F. Supp. 3d 762, 783-84 (E.D.N.Y. 2019) (concluding that concerns about error rate, the circular sufficient agreement standard, and lack of

¹⁵ In *Adams*, 444 F. Supp. 3d at 122-56, the government actually withdrew its intention to offer the examiner's opinion that shell casings were a match to the recovered handgun, and instead sought to offer only what the Court ultimately permitted as to his observations, not a comparison-based opinion.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 acceptance in scientific community do not permit government to demonstrate sufficient
2 reliability of source attribution to specific firearm recovered, and limiting testimony
accordingly).

3 This Court should join the chorus, not because there is strength in numbers, or for
4 conformity's sake, but because the conclusions and judgments limiting examiner testimony
5 are eminently sound, fair and in line with research developments and the weight of the
6 scientific community. As it concerns firearms identification evidence, unqualified admission
7 was the past, limitation seems to be the present, and exclusion appears to be the future.
8

9 **C. Black-box Studies Cast More Doubt**

10 Much of the discussion in the PCAST Report and elsewhere refers to “black-box”
11 studies, of which there are less than a handful. The first one was the 2014 Baldwin/Ames
12 Lab study, referenced *supra*, and attached as Exhibit E. There, following an “open-set”
13 design that more closely resembles actual casework, the error rate as to false positives (a
14 claimed match where it was known that the ammunition component had not been fired
15 from that weapon) was estimated at 1 in 66, or about 1.51%, although reasonably as high as
16 2.2%. Exhibit C, at 110-11; Exhibit E, at 15-17. While such an error rate might not appear
17 terribly high, it is still troubling that potentially 1 of every 66 (or as low as 46) defendants
18 convicted on the basis of such evidence, might be innocent. And it also casts doubt on
19 much lower error rates in closed-set studies that allow for inferential deductions that
20 casework does not. Critically, and as discussed more below, when inconclusive responses
21 are not excluded from the error rate, it rises to nearly 35% (22 false positives added to 735
22 inconclusives, divided by 2,178 different source comparisons). *See Tibbs*, 2019 D.C. Super.
23
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 Ct. LEXIS 9, at *57-63 (discussing problem of inconclusives in Ames study and others);
2 *Adams*, 444 F. Supp. 3d at 1265 (noting that the lack of negative consequence for an
3 “inconclusive,” incentivizes test subjects to record lower false positive rate than in
4 casework). The frequency of inconclusive responses¹⁶ (about one third) would appear to
5 cast doubt on the premise that examiners can reliably discern individualized toolmarks and
6 make source attributions. Moreover, the Ames study was not published in a peer-reviewed
7 journal, and is available instead on the AFTE website.
8

9 The results of a second Ames Labs/FBI study were released in October 2020,
10 although also not in a peer-reviewed publication. See Stanley A. Bajic et al., *Report: Validation*
11 *Study of the Accuracy, Repeatability, and Reproducibility of Firearm Comparisons*, Oct. 7, 2020
12 [attached as Exhibit H]. This is where things get interesting. Exhibit D, at 55 (describing
13 how the report “backfired on proponents . . . and quite unintendedly presents a scathing
14 indictment of firearms identification methodological reliability”). Ostensibly the study finds
15 an error rate as to false positives (declaring a match or identification as to a known non-
16 match) of about .7% for bullets, and about .92% for cartridge casings. Exhibit H, at 34-36.
17 However, these error rates ought to be taken with a huge grain of salt, as responses of
18
19
20
21

22
23 ¹⁶ Baldwin’s accompanying report appears to discount the possibility that the inconclusives
24 resulted simply from insufficient markings on the samples: “The [large] fraction of samples reported as
25 inconclusive cannot be attributed to a large fraction of poorly marked knowns or questioned samples in this
group.” Exhibit E, at 19. In other words, the problem was not with the samples.
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 “inconclusive” were, again, not counted as “error” or included in the denominator when
2 tabulating error rate. This is highly problematic.

3 In casework, an “inconclusive” is less of a problem, as it is unlikely to lead to false
4 convictions or be incriminating, and where it is unknown whether the casing or bullet was
5 fired by a particular firearm. But in validation studies, the ground truth (whether a bullet or
6 casing was fired by a particular gun) *is* known. Exhibit D, at 56. That is the whole point.
7 Thus, an “inconclusive” is, objectively, an incorrect answer where only one of two
8 possibilities exist, match or elimination. Permitting respondents to answer in that fashion
9 and excluding inconclusives from the error rate “allows a respondent to pick and choose
10 which questions he/she feels most confident in answering.” Exhibit D, at 56. Another
11 highly respected voice at the intersection of law and science finds this “absurd as a research
12 design” and a profound mistake to fail to include inconclusives when calculating error rates.
13 *Id.*; see also I.E. Dror, N. Scurich, *(Mis)use of Scientific Measurements in Forensic Science*, 2 Forensic
14 Science International: Synergy, 333, 337 (Aug. 2020) [attached as Exhibit I] (“[E]rroneous
15 inconclusives, and inconclusive evidence should be accounted for in error rate studies, and
16 failing to do so results in measurements that are misleading and inaccurate”).

17 Therefore, when “inconclusives” are not excluded as incorrect answers, the error
18 rates approach or exceed the probability of calling heads or tails in a random coin toss. For
19 bullets, the error rate is an astounding 53% (worse than flipping a coin), whereas the rate for
20 cartridge casings rises to 44%. See Exhibit D, at 57. Additionally, for both claimed matches
21 and eliminations, Tobin and colleagues present data from Ames II as to troublingly low rates
22
23
24
25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 of repeatability (the same examiner coming to the same prior opinion) and reproducibility
2 (different examiners coming to the same conclusion as a previous examiner), the two key
3 metrics of scientific validation, and not what one would expect or require for a discipline
4 featuring reliable methods and that is reliably applied. Exhibit D, at 57. Also significant is
5 an apparent survivorship bias, where *only about 31% of initial respondents* (who were self-
6 selected in the first place and knew they were being tested, presenting serious problems in its
7 own right¹⁷) *actually completed the study*. See Exhibit D, at 56-58; Exhibit H, at 15 (noting that
8 of the initial 256 participants, only 80 had finished all six mailings of test-packet analyses and
9 that a “high dropout rate has been considered a warning flag for possible test bias”). It is
10 “quite likely that the rates of error and poor repeatability and reproducibility rates would be
11 even worse (1) for the general population of examiners, and (2) that those dropping out of
12 the study could well be those most likely to make errors in the crush of heavy workloads[.]”
13 Exhibit D, at 58.

14
15
16 A widely-cited study published in the AFTE Journal in 2018, *see* Mark A. Keisler et
17 al., *Isolated Pairs Research Study*, 50 AFTE JOURNAL 56, 57 (2018) [attached as Exhibit J], is
18 similarly flawed: “Inconclusive answers are not considered incorrect.” This is inconsistent
19 with the goal of providing samples that simply either matched or did not. Exhibit J, at 56-
20

21
22
23
24 ¹⁷ Tobin notes that because of self-selection, “only examiners most confident in their abilities
25 volunteered as respondents. They likely do not constitute representative sampling of the firearms
identification field at large as to confidence, competence, and skillsets.” Exhibit D, at 56.
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

57. Of 1,008 true exclusions possible, subjects made only 805 exclusions, which, by one metric, can be characterized as an error rate around 20%.

Its reported error rate of 0% as to false positives is problematic, both as compared to other studies, and its apparent attribution to its human test-subjects the extremely inhuman trait of infallibility. Further, evidently there was no testing as to the difficulty of the task, i.e., whether it presented the sort of challenges examiners might face in the field. Exhibit J, at 58 (“[I]t is unclear if participants used quality assurance measures, such as verifications, when conducting the research.”). The study also did not make clear whether the subjects conformed to or deviated from their respective laboratory policies, or to how they were selected for participation in the first place. *Id.* Nor was the study peer reviewed by anyone outside of the narrow AFTE community so as to permit more meaningful, independent scrutiny.

Most studies were seemingly designed and administered to “validate” pre-existing belief about the accuracy of firearms identification opinion. *See, e.g.*, Exhibit J. They were not employed to falsify a given hypothesis, a key tenet, perhaps *the* key feature of the scientific approach¹⁸. This is one reason among many that the general scientific community, as seen, for example in the NRC and PCAST reports, has expressed doubt as to such studies and the minimal error rates they report, and that legitimate, science-based research has not

¹⁸ From a scientific standpoint, it is problematic to design and administer a study for the purpose of attempting to “validate” a claimed theory. That is not the way things are supposed to work, at least if the goal is testing the validity of a given hypothesis.

1 confirmed its foundational validity, especially as to discernible uniqueness or specific source
2 attribution. *See* Clifford Spiegelman & William A. Tobin, *Analysis of experiments in forensic*
3 *firearms/toolmarks practice offered as support for low rates of practice error and claims of inferential*
4 *certainty*, 12 LAW, PROBABILITY & RISK 115, 131 (2012) [attached as Exhibit D-3]
5 (concluding, following a thorough review of available evidence and studies, that a “statistical
6 foundation supporting inferences of specific source attribution for firearms/toolmarks does
7 not exist”).

8
9 And courts have taken notice. *Adams*, 444 F. Supp. 3d at 1266-67 (concluding that
10 it is “not clear that that those results [claiming firearms identification] are the product of
11 a *scientific* inquiry, [and that nothing explains] how or why [the government’s expert] reached
12 his conclusion in any quantifiable, replicable way.”); *Shipp*, 422 F. Supp. 2d at 762 (limiting
13 expert testimony and precluding opinion as to a match “will prevent the jury from placing
14 unwarranted faith in an identification conclusion based on the AFTE Theory, which the
15 current research has yet to show can reliably determine” source attribution); *Ross*, 68 Misc.
16 3d at 918 (excluding opinion testimony on the “significance of any marks other than class
17 characteristics, as the reliability of that practice in the relevant scientific community as a
18 whole has not been established. Moreover, any opinion based in unproven science and
19 expressed in subjective terms such as ‘sufficient agreement’ or ‘consistent with’ may mislead
20 the jury and will not be permitted.”). In short, because of myriad flaws, even the better-
21 designed studies still raise as many questions as to validity and reliability as the answers they
22 provide. At best, they fail to offer convincing support for the reliability of firearms
23
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 identification. And at worst, they raise the prospect that firearms identification, when it
2 comes to individualized source attribution, is entirely unreliable.

3 **D. The Declaration of William A. Tobin Casts Even More Doubt**

4 In connection with this motion and the requested *Daubert* hearing, Mr. Cloud
5 engaged the services of an expert, Bill Tobin, to opine on firearms identification's
6 foundational validity. Mr. Tobin – a trained metallurgist and materials scientist with many
7 years of service with the FBI before his retirement and entry into a consulting practice – is
8 one of the most respected voices in the scientific and legal communities as it concerns
9 toolmarks and firearms identification. His CV is attached as Exhibit D-1. The breadth and
10 depth of his expertise is likely unmatched. Mr. Tobin has prepared a lengthy declaration,
11 attached as Exhibit D. We do not attempt to summarize all of Mr. Tobin's opinions and
12 conclusions and qualifications; they are too detailed and numerous, and defy such an effort.
13
14

15 From a technical, scientific and empirical standpoint, Mr. Tobin paints a worrisome
16 picture as to the state of the reliability and foundational validity of firearms identification,
17 particularly as to the individualized level of source attribution and in the aftermath of the
18 second Ames black-box study. He notes the critical absence of scientific criteria – to include
19 falsifiability, repeatability and reproducibility, among others - which leaves the practice
20 entirely subjective, and thus not scientific. Exhibit D, at 17-25; 41-42. Mr. Tobin further
21 describes an absence of evidence as to the uniqueness of individualized markings on
22 cartridge components, and that any uniqueness (if it exists) would be discernible to the
23 examiner. The premises necessary for source attribution are, therefore, unfounded and
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

unproven. In particular, he finds no evidence that firearms examiners can reliably discern subclass characteristics (from the manufacturing process) from purportedly individual characteristics, another key premise on which individualized source attribution necessarily depends. Exhibit D, at 23, 28-31, 43. He finds unacceptably low rates of repeatability and reproducibility, essential indicia of reliability, and that the vagueness and circularity inherent in the AFTE theory of identification establish only subjective criteria, providing “no scientifically acceptable guidance whatsoever.” Exhibit D, at 17.

Mr. Tobin also notes an elevated risk of misattribution when it comes to impressions made by firing pins and breech face markings on cartridge cases¹⁹, as opposed to striae on bullet fragments:

[I]dentifications made with firing pin impressions are known in the firearms identification community to be one of the least reliable components on which to render identifications, and breech face identifications are particularly vulnerable to error because the characteristics imparted to cartridge cases are almost entirely subclass from the manufacturing processes used to form the breech face. Both components, firing pins and breech faces, will transfer characteristics that would be confusingly similar with their respective components in different firearms, and virtually analytically indistinguishable in an inductive inferential test environment (case work) from the other firearms in the same production lot.

Exhibit D, at 15-16; *see also id.* at 23, 34-35, 44-45. To sum up, Mr. Tobin concludes that:

firearms identification examiners themselves have inadvertently demonstrated that they cannot do what they claim to be able to do: render opinions of individualization with any acceptable degree of accuracy or reliability. The only scientifically and forensically defensible opinion for the current state of firearms identification is that, *in the opinion of the examiner*, a specific firearm could not be

¹⁹ This is important here because what both Drury and Van Arsdale purport to match to each other or to the Ruger rifle at issue, are spent cartridge cases, not bullets or bullet fragments.

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 eliminated as the firing platform for evidentiary bullets or cartridge cases,
2 tantamount to allowing only class characteristics to be the basis for the opinion.

3 Exhibit D, at 58 (emphasis in original).

4 Of course, whether expert testimony is scientific or not is not necessarily
5 determinative of whether it is admissible, and to what extent. Courts sometimes admit the
6 testimony of experts who do not practice a scientific discipline, but who possess other
7 technical or specialized knowledge. *See United States v. Hankey*, 203 F.3d 1160, 1168 (9th Cir.
8 2000). But whether or not the field and methodology is scientific, the key point, as far as
9 *Daubert* is concerned, is that it is reliable and characterized by underlying validity. Science or
10 the scientific method are, in some respects, a proxy for reliability, a method for assuring
11 those ends. And the point, whether or not firearms identification (especially to the level of
12 claimed match or specific source-attribution) is scientific or science-based, is that it lacks
13 reliability and validity.
14

15 **VI. The *Daubert* Factors Militate Against Admission of Firearms**
16 **Identification Testimony and a Finding that the Practice is Reliable**

17 Moreso than for other disciplines, the relevant *Daubert* factors tend to bleed into
18 one rather than remaining analytically distinct. While we attempt to discuss each factor
19 separately, they are at least somewhat inter-related, and the discussion reflects it. Or perhaps
20 this is simply an indication that the reliability of firearm identification evidence is severely
21 wanting. *See Adams*, 444 F. Supp. 3d at 1259 (recognizing that “the flaws in this kind of
22 quasi-scientific methodology are more fundamental than just the enigmatic process by which
23 the ultimate conclusion is reached, and those flaws end up being carried over into every
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 other factor.”). Either way, each *Daubert* factor is neutral, or weighs in favor of exclusion of
 2 source attribution statements, and each raises significant concerns as to the reliability of
 3 firearms identification practices.

4 **A. Whether the Theory or Technique has been Tested**

5 While firearms identification and AFTE theory has been tested to some extent, the
 6 worth or merit of those studies is either in dispute (at best), or is not sufficient to
 7 demonstrate the reliability of the technique of matching bullets or shell casings to an
 8 individual firearm, beyond class characteristics. *See Shipp*, 422 F. Supp. 2d at 776
 9 (recognizing that “courts should avoid placing too much trust in studies that may
 10 overestimate the method’s reliability”). The PCAST Report, to the point, lamented the lack
 11 of Black-box type studies leading up to 2016, and the years since have featured similarly
 12 flawed or limited reports. Exhibit C, at 106-09. Mr. Tobin too, for example, finds “no
 13 comprehensive or meaningful empirical study that constitutes acceptable scientific validation
 14 of the forensic practice of firearms identification” and details the problems in study design
 15 that do not mirror inductive casework. Exhibit D, at 52. And the reports that have
 16 surfaced recently, e.g., Ames II [Exhibit H], have provided more reason to doubt than
 17 validate firearms identification testimony. *See supra*; Exhibit D, at 55-58.

18 Unless testing can be shown to yield reliable results, it, by itself, does not move the
 19 needle in favor admissibility. Testing just for the sake of testing is not part of the *Daubert*
 20 analysis. *See* Fed. R. Evid. 702 adv. committee notes (describing this factor as “whether the
 21 expert's theory can be challenged in some objective sense, or whether it is instead simply a
 22 Motion to Exlude or Alternatively Limit Toolmark Identification Evidence

1 subjective, conclusory approach that cannot be reasonably assessed for reliability”). The
2 more salient question is precisely *what is being tested?* Because of the AFTE theory’s inherent
3 subjectivity and circular, tautological reasoning, it does not lend itself well to empirical
4 testing.

5 Courts have looked at two indicators here, both of which are problematic. *See, e.g.,*
6 *Adams*, 444 F. Supp. 3d at 1257 (describing that testability, as *Daubert* factor, refers to
7 whether a methodology is falsifiable and replicable, and distinguishes science from other
8 disciplines). With regard to falsifiability, the practice has “no scientifically acceptable
9 protocol” in that it lacks a refutable hypothesis. Exhibit D, at 17, 19-20. This is because,
10 beyond class characteristics, it depends on the assumption that all firearms leave unique,
11 individualized markings to the exclusion of all others, which itself has not been proven. *See*
12 Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*,
13 61 VAND. L. REV. 199, 218 (2008) (“Forensic scientists are not able to link . . . a toolmark .
14 . . . to its unique source, but they assert that ability every day in court.”). And while this might
15 sound like it sets the bar unreasonably high, it is the bar that science itself demands before
16 being recognized as such.

17 As to replicability, this is constrained by the subjective nature of the examiner’s
18 opinion in finding sufficient agreement, as well as his/her training and experience, and the
19 circular, tautological way in which he or she arrives at that conclusion. It is not surprising
20 then, that both repeatability and reproducibility fare poorly in the recent Ames study.
21 Exhibit D, at 55-57. “Not only is the AFTE method not replicable for an outsider to the
22 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
23
24
25

1 method, but it is not replicable between trained members of AFTE who are using the same
 2 means of testing.” *Adams*, 444 F. Supp. 2d at 1263. This is because it “cannot be explained
 3 in a way that would allow an uninitiated person to perform the same test in the same way[.]”
 4 *Id.* Thus, on balance, the limited value of testing thus far does not militate in favor of
 5 admission. *See id.* at 1263 (concluding the AFTE method is “not replicable – and not
 6 testable” and that that factor “weighs heavily against admissibility under *Daubert*”).

8 **B. Whether the Theory or Technique has been Subjected to**
 9 **Peer Review and Publication**

10 Firearms identification as to source attribution does not do much better when it
 11 comes to peer review. This factor too weighs in favor of exclusion, or at the least, limitation
 12 for many of the same reasons. The AFTE has a vetting process, but whatever nominal peer
 13 review and publication has occurred generally, is overcome or outweighed by the NRC and
 14 PCAST Reports, or by the combined force of those analyses.

15 Whether firearms identification has been adequately subjected to peer review
 16 depends on what is considered a peer. If by peer, we are talking about other members of the
 17 AFTE community, then there is some support for the notion that AFTE theory and
 18 scholarship is subject to peer review. However, Mr. Tobin describes the “*AFTE Journal*, [as]
 19 realistically the sole trade journal with virtually no scientifically acceptable external peer
 20 review or even access by scientific scholars or other non-AFTE members for the many
 21 decades of its application and admissibility.” Exhibit D, at 16. Further, the AFTE’s peer-
 22 review process is open, such that the author and reviewer know the identities of the other,
 23 which does not remove personal and institutional biases. *See Tibbs*, 2019 D.C. Super. Ct.
 24 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
 25

1 LEXIS 9, at *30 (finding that open review “diminishes the extent to which proponents of
2 firearms and toolmark identification evidence can claim that its articles have been subjected
3 to meaningful, stringent peer review”). The AFTE Journal is also closely guarded, and this
4 “restricted access effectively forecloses the type of review of the journal's publications by a
5 wider community of scientists, academics, and other interested parties that could serve as an
6 important mechanism for quality assurance.” *Id.* at *31. In this way, AFTE theory “avoids
7 the scrutiny of scientists and academics outside the field of firearms and toolmark analysis.
8 These limitations significantly diminish the stringency of the review that a study published in
9 the *AFTE Journal* can be said to have undergone, even after its publication.” *Id.* at *32.
10 Thus, as is the case when defining the parameters of the relevant scientific community, *infra*,
11 peer review here is generally characterized by AFTE insularity or opacity, and not genuine
12 peer review as practiced in scientific disciplines.
13

14
15 For these reasons, as well as others, *Tibbs* severely discounted the value of the
16 AFTE’s peer review process, noting that reviewers, all AFTE adherents, “have a vested,
17 career-based interest in publishing studies that validate their own field and methodologies.”
18 *Id.* at *33. What are they supposed to say? Children might unanimously agree that it is a
19 good idea to permit kids to eat as many cookies as they want, but such a “peer review”
20 process hardly denotes the validity of such position from a science or health-based
21 perspective. Validation by true believers does not necessarily satisfy the strictures of peer
22 review, or confer any broader reliability. And that multiple adherents and practitioners of
23 the same AFTE theory might render the same conclusions is only as meaningful as the
24
25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 underlying theory²⁰; if the premise is flawed and without foundational validity, the rest is not
2 terribly significant.

3 Moreover, the peer review community, to the extent it is composed of other
4 examiners versed in the AFTE Theory of Identification, “may be trained and experienced in
5 the field of firearms and toolmarks identification, but not necessarily have any specialized or
6 even relevant training in research design and methodology.” *Tibbs*, at *32-33. From a
7 scientific or empirical standpoint, this is a problem. When the review process is performed
8 by scientists or researchers well versed in those methods, the results are hardly similar. The
9 PCAST and NRC Reports are far more skeptical than the AFTE adherents and practitioners
10 who publish in the organization’s journal. For example, the PCAST Report takes issue with
11 the paucity of black-box studies compared to closed-set type designs. Exhibit C, at 106-11.
12

13 Curiously, however, of the three studies that might be characterized as black-box-
14 type research designs, discussed *supra* at 22-26, neither of the Ames studies were published in
15 a peer-reviewed journal. “Unfortunately, the only studies with the more appropriate design
16 for assessing reliability—the [first] Ames Laboratory study and the Keisler study [Exhibits E
17 & J] —have not, as described *supra*, undergone meaningful, independent peer review prior to
18 publication.” *Tibbs*, 2019 D.C. Super. Ct. LEXIS 9, at *56. Therefore, whatever peer review
19 process AFTE employs would appear to have limited value and structural flaws, and does
20
21

22
23
24 ²⁰ Peer review is also problematic in this area because of the subjective nature of standards and
25 what constitutes sufficient agreement.

1 not bestow reliability. *See id.* at *35 (“the *AFTE Journal*’s use of reviewers exclusively from
2 within the field to review articles created for and by other practitioners in the field greatly
3 reduces its value as a scientific publication, especially when considered in conjunction with
4 the general lack of access to the journal for the broader academic and scientific community
5 as well as its use of an open review process”).

6
7 **C. Existence and Maintenance of Standards that**
8 **Control the Technique’s Operation**

9 This factor, because of its subjectivity inherent in the AFTE theory of
10 identification, is the one that has caused courts the most problems in terms of assessing the
11 reliability of firearms identification evidence. The problem is not that the technique or
12 discipline is completely without standards. Instead, it is that whatever standards exist are
13 largely subjective or intrinsic to the examiner, and not objective. *See, e.g., Shipp*, 422 F. Supp.
14 2d at 780 (critiquing the “near total subjectivity countenanced by the AFTE Theory, where
15 there is no actual guidance for what comprises ‘sufficient agreement’”). Even that would
16 not necessarily be determinative as to reliability or admissibility. If subjective standards
17 could be more objectively or empirically tested, that would help. But the studies, as
18 addressed elsewhere, are often inherently flawed, and the ones that feature fewer flaws
19 hardly demonstrate foundational validity. As such, the standards remain largely opaque.

20
21 By its own terms, the AFTE theory relies on subjective standards. This is not
22 seriously in dispute. But recently, courts have increasingly targeted the circular, tautological
23
24

reasoning²¹ that leads an examiner to find sufficient agreement, on the basis that the purported standard does not “*mean anything.*” *See Adams*, 444 F. Supp. 2d at 1262; *see also Glynn*, 578 F. Supp. 2d at 572 (describing theory of identification as “inherently vague”); *Green*, 405 F. Supp. 2d at 114 (describing AFTE theory as “either tautological or wholly subjective”). Or it means whatever the examiner decides, without observable, measurable or reproducible guideposts, and is too dependent on training and experience. *See Tibbs*, 2019 D.C. Super. LEXIS 9, at *71-72 (“Unlike the standards underlying opinions in other fields, the AFTE theory provides no objective yardstick to support or explicate the expert's opinion; instead, the expert is left to rely on her own thoughts and conclusions based only on the vagaries of her own training and experience” [and “relies entirely on subjective judgment”].) As Judge Garaufis put it, “the subjectivity of the AFTE Theory raises serious

²¹ *Tibbs* gets to the heart of this circularity, and adroitly recognizes that AFTE theory:

can only be characterized as entirely tautological: an opinion of common origin can be rendered when the surfaces of the two examined items are in “sufficient agreement,” which exists not when some objective measure is satisfied, but when the examiner determines, based on her training and experience, that it would be a “practical impossibility” for the two items not to share a common origin. In other words, this protocol permits the ultimate finding of “sufficient agreement” whenever an individual examiner concludes that she would be hard pressed (for reasons not specified in the governing standard) to find such similar markings on casings or bullets fired by different firearms. Although AFTE has attempted to use terms like “sufficient agreement” to resemble terminology that one would find in an objective or scientific standard, in the end it simply leaves the determination of common origin to the standardless, undefined judgment of an individual examiner. Therefore, under this so-called standard, the process for determining what constitutes a “match” lacks defined criteria; it is merely unconstrained subjectivity masquerading as objectivity.

Tibbs, 2019 D.C. Super. LEXIS 9, at *68-69.

1 concerns about the reproducibility of examination results across labs and examiners and the
2 accuracy of those results. Therefore, the court finds that the subjective and circular nature of
3 AFTE Theory weighs against finding that a firearms examiner can reliably identify when two
4 bullets or shell casings were fired from the same gun.” *Shipp*, 422 F. Supp. 2d at 782.

5
6 These concerns as to the lack of objective standards and the impact on scientific
7 validity echo those previously raised by scholars and scientists, including Mr. Tobin. *See*
8 William A. Tobin & Peter J. Blau, *Hypothesis Testing of the Critical Underlying Premise of Discernible*
9 *Uniqueness in Firearms-Toolmarks Forensic Practice*, 53 JURIMETRICS J. 121, 127 (2013) [attached
10 as Exhibit D-2] (“Because the [AFTE Theory] is comprised of such vague and subjective
11 terms, with no underlying protocol, it does not incorporate, or even allow for, two critical
12 cornerstones of true scientific endeavor: repeatability and reproducibility”); Spiegelman &
13 Tobin, *Analysis of Experiments, supra*, [Exhibit D-3], at 131 (criticizing lack of meaningful,
14 comprehensive standard operating procedures and concluding that “[u]ntil the
15 firearm/toolmark community has detailed SOPs, the field will remain without the key tool
16 needed for a scientific procedure to exist”).

17
18 In Mr. Tobin’s more recent declaration, he expounds on these ideas, noting that
19 the AFTE Theory of Identification:

20 espouses exactly what the Scientific Method attempts to eliminate (subjectivity)
21 . . . The Scientific Method was designed to obviate subjective influences; the
22 practice of firearms identification is entirely (100%) subjective once an
23 appropriate sample pool is identified for comparison, the most critical portion
24 of the methodology. Further, just what constitutes “sufficient agreement” is not
25 defined and is left up to the individual examiner’s subjective interpretation and

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 will vary between and among examiners depending on experience, regional
2 practice, and other influential variables.

3 Exhibit D, at 17 (omitting footnote in original). Put more concretely, the examiner's ability
4 to discern sub-class from individual characteristics depends purely on subjectivity. Exhibit
5 D, at 29-30, 40-42; *Shipp*, 422 F. Supp. 2d at 781. For these reasons, the lack of objective
6 standards present an unsurpassable roadblock to reliability and "strongly militates against the
7 admission of expert witness testimony in the field of firearms and toolmark analysis." *Tibbs*,
8 2019 D.C. Super. LEXIS 9, at *71-72.

9 **D. Error Rates**

10 Much of the focus in toolmark challenges and court decisions has been on error
11 rates, known or potential. And for good reason. Dean Faigman referred to error rates as,
12 effectively, the "cornerstone of foundational validity." *Ross*, 68 Misc. 3d at 910 (quoting
13 testimony from motions hearing). It is difficult to believe that a given method with an
14 unacceptably high error rate would pass muster under *Daubert*, even if it were somehow able
15 to satisfy the other factors. Discussing error rates as to toolmark identification is something
16 of a *Rorschach* test; there is something for everyone to grasp on to. But on balance, it should
17 be cause for serious concern.
18

19 The significance of known [or potential] rates of error is, of course, constrained by
20 the quality, design and methodology of the studies that produce them. *See e.g., Tibbs*, 2019
21 D.C. Super. LEXIS 9, at *56 ("Because of these significant limitations of the closed and set-
22 based studies, the vast majority of studies that the field relies upon to establish its
23 foundational validity simply do not provide an adequate basis to do so."). In other words,
24 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
25

1 the claimed error rate is only as good as the study that produced it, and what it was actually
2 measuring. The PCAST Report, for example, “cautioned that not all study designs are
3 equally trustworthy when predicting error rates.” *Shipp*, 422 F. Supp. 2d at 777.

4 *Tibbs* identified three problems, as far as error rates are concerned:

5 Most of the studies suffer from basic, threshold design flaws that undermine the
6 value of their stated results. Second, the reliance of most of these studies on
7 “closed” and/or “set-based” design structures substantially limit the reliability of
8 the error rates claimed in these studies. Third, and perhaps most significantly, the
9 studies permit participants to label toolmark comparisons as “inconclusive”
without adequately assessing the impact of such inconclusive determinations on the
results of the study as a whole.

10 2019 D.C. Super. LEXIS 9, at *40-41. The studies that purport to feature minimal rates of
11 error, at 1% or less, tend to be the sort of “closed set” designs that the scientific community
12 eschews: “The claims of ‘near zero’ or ‘less than 1 percent’ rates of error virtually
13 unanimously declared by forensic firearms examiners are vacuous, comprise *ipse dixit*, and
14 are without scientific foundation. The only basis for the claims derives from ‘closed-set’
15 proficiency tests that entail myriad flawed aspects in methodology and do not mirror actual
16 casework.” Exhibit D, at 48.

17 Things start to look quite different, and the error rates rise, sometimes significantly,
18 when study designs require the inductive process used in the field on a given lab
19 examination. For example, an open-set design that more closely resembled actual casework
20 found an error rate (for false positives) at or above 2%, and “roughly 100-fold higher” than
21 closed set studies. Exhibit C, at 109; *see also Adams*, 444 F. Supp. 3d at 1264 (noting that
22 “higher error rates tend to arise from the studies that most closely resemble the real-world
23 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
24
25

1 conditions of toolmark testing. The lowest rates arise from the “closed-set” tests). From a
2 reliability standpoint, this is concerning at the least. *See Shipp*, 422 F. Supp. 3d at 778
3 (finding, on the basis of the “study that most closely resembled fieldwork [Ames I],” that the
4 error rate “cautions against the reliability of the AFTE Theory”).

5
6 Secondly, as discussed above at 23-25, low error rates have resulted from
7 calculations that tend not to treat “inconclusive” responses as error at all. The notion of
8 counting only a false positive or false negative as error thus minimizes the claimed error rate.
9 This is a problem, given that in a study there is either a “match” [identification] or a non-
10 match [exclusion]. Where the objective or ground truth only permits two possibilities
11 (match or non-match), an inconclusive is “error” and certainly relevant to the validity of the
12 findings and the reliability of the discipline. *See Tibbs*, 2019 D.C. Super. LEXIS 9, at *60-62
13 (expressing the view that “inconclusives” should be treated as errors of some kind, even if
14 not akin to false positives, because there should be sufficient information for either
15 identification or exclusion); Itiel E. Dror & Glenn Langenburg, *“Cannot Decide”: The Fine Line*
16 *Between Appropriate Inconclusive Determinations Versus Unjustifiably Deciding Not to Decide*, 64 J.
17 *FORENSIC SCI.* 10, 13 (2019) (“An inconclusive determination is an erroneous decision
18 because the evidence does not support that decision”). Therefore, we must consider the
19 effect inconclusives have on the study overall, and its resulting error rate. As discussed
20 above, at 24-25, the error rates that account properly for inconclusives, for example, in the
21 second Ames study, are extraordinarily high.
22
23
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 The large number of inconclusives in these studies is problematic even apart from
2 the effect on the error rate, and “greatly reduces their persuasive force in establishing the
3 ability of a firearms and toolmark examiner to make accurate source determinations.” *Tibbs*,
4 2019 D.C. Super. Ct. LEXIS 9, at *63. If firearms identification is as reliable as its
5 proponents claim, then why are its practitioners so frequently unable to reach decisions in
6 testing that is designed to enable them to reach an opinion of identification or exclusion?
7 The answers to that question are unsatisfactory. One is that the fact that the examiners are
8 being tested, which results in more inconclusives than does examination in the real world.
9 *See* Dror & Scurich, *supra* [Exhibit I], at 336 (“Examiners resort to making more inconclusive
10 decisions during error rate studies than they do in casework”). But if examiners are more
11 cautious or hesitant during studies, this means that they are more apt to make errors,
12 including false positives, during actual casework, and that the error rate is likely higher than
13 as reported in the studies (even if inconclusives are *not* treated as incorrect responses). For
14 all these reasons, the error rate, known or potential, “currently weighs against the admission
15 of source attribution statements made by a firearms and toolmark examiner”. *Tibbs*, 2019
16 D.C. Super. LEXIS 9, at *65.

19 **E. General Acceptance Within the Relevant Community**

20 This factor again raises many similar concerns as the discussion on peer review.
21 How we frame the relevant community makes a difference, and *Kumbo Tire* cautions courts
22 against placing parameters that are too restrictive. 526 U.S. at 151. Here, there is wisdom in
23 narrowly limiting that community to firearms examiners or AFTE adherents.

24 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
25

1 If we are speaking solely of firearms examiners or AFTE members, then its
2 theory of identification, and individualized source attribution has achieved general
3 acceptance. *See Shipp*, 422 F.3d at 782 (describing how “courts have, in cursory fashion,
4 identified toolmark examiners as the relevant community, and have summarily determined
5 that AFTE theory is generally accepted in that community”); *Tibbs*, 2019 D. C. Super.
6 LEXIS 9, at *73 (noting that courts finding that AFTE theory has achieved general
7 acceptance “have generally limited the scope of the so-called ‘relevant community’ to the
8 specific community of firearms and toolmark examiners”). This is the crux of the problem,
9 and is not surprising: the community that performs peer review and expresses general
10 acceptance is composed almost entirely of true believers, or AFTE brethren. Astrologers
11 believe in the legitimacy of astrology. Diviners believe wholeheartedly in divination.
12 Necromancers believe in raising the dead²². And toolmark analysts believe in the reliability
13 of firearms identification; their livelihoods depend on it.
14
15

16
17
18 ²² In evoking such examples, defense counsel in no way minimizes the seriousness of the instant
19 charges, or of the fact that five people lost their lives as a result of the charged conduct. But that is the
20 point: the seriousness of the charges (and of the penalties, if proven) requires evidence that is
21 correspondingly serious and reliable. And toolmark identification does not meet that standard. More, the
22 analogies come straight from seminal pronouncements in *Kumho Tire*, 526 U.S. at 151: “Nor, on the other
23 hand, does the presence of *Daubert*’s general acceptance factor help show that an expert’s testimony is
24 reliable where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called
25 generally accepted principles of astrology or necromancy.”

22 Further, no less a scholar than Judge Jed. S. Rakoff of the Southern District of New York also
23 employed the same analogy: “[J]ust as astrologers will attest to the reliability of astrology, defining ‘peer’ in
24 terms of those who make their living through handwriting analysis would render this *Daubert* factor a
25 charade.” *Almeciga, v. Center for Investigative Reporting, Inc.*, 185 F. Supp. 3d 401, 420 (S.D.N.Y. 2016). As
Upton Sinclair famously said, “It is difficult to get a man to understand something, when his salary depends
on his not understanding it.” *I, Candidate for Governor: And How I Got Licked* (1935).

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 But the imprimatur of the narrow peer community alone, in this context at least,
2 does not confer upon the practice any degree of reliability or scientific validity, and may even
3 be counterproductive. *See Adams*, 444 F. Supp. 2d at 1266 (suggesting that “the widespread
4 acceptance within the law enforcement community may have created a feedback loop that
5 has inhibited the AFTE method from being further developed”); *Ross*, 68 Misc. 3d at 913
6 (“The targeted use of AFTE theory by law enforcement investigators, under pressure and
7 with potential for confirmation bias, limits the degree of intellectual rigor and detachment
8 that counts as neutral scientific expertise”) (citing *Shipp*). Astrology is not reliable simply
9 because other astrologers vouch for it; indeed, in spite of such vouching, it might not be
10 reliable at all.

12 Accordingly, when discussing general acceptance, judges have noted the concerns
13 of broader scientific community, as expressed in the PCAST Report and elsewhere, and have
14 not limited the “relevant community” to just firearms examiners. *See, e.g., Adams*, 444 F.
15 Supp. 2d at 1266 (“[W]here the scientific community at large disavows the theory because it
16 does not meet the parameters of science, I cannot find that the AFTE method enjoys
17 ‘general acceptance’ in the scientific community”); *Shipp*, 422 F. Supp. 3d at 782-83 (agreeing
18 that relevant community extends beyond AFTE members to “pre-eminent scientists and
19 scholars” who authored PCAST and NRC reports); *Tibbs*, 2019 D.C. Super. LEXIS 9, at *73
20 (expressing the view that if *Daubert* and Rule 702 “are to have any meaning at all, courts
21 must not confine the relevant scientific community to the specific group of practitioners
22 dedicated to the validity of the theory—in other words, to those whose professional
23 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence
24
25

1 standing and financial livelihoods depend on the challenged discipline”); *Ross*, 68 Misc. 3d at
2 914 (“[A] number of district courts hearing *Daubert* challenges have broadened the
3 relevant scientific community to take into account the contrasting views of mainstream
4 researchers. Consequently, the scope of permissible expert toolmark testimony is narrowing
5 overall.”).

6
7 When we take into account the views of mainstream scientists, firearms
8 identification, at least to the level the government proposes with Van Arsdale’s source
9 attribution testimony, falls well short of general acceptance. *See Shipp*, 422, F. Supp. 2d at
10 782-83 (concluding that “AFTE Theory has not achieved general acceptance in the relevant
11 community”); *Tibbs*, 2019 D.C. Super. LEXIS 9, at *74 (“The conclusions of the NRC and
12 PCAST reports indicate that the wider academic and scientific community does not
13 necessarily generally accept this theory [and] the government failed to show general
14 acceptance outside of the field of firearms and toolmark practitioners”); *Ross*, 68 Misc. 3d at
15 915 (“This broad community has no consensus on the more subjective findings and
16 conclusions of toolmark examiners. The vague ‘sufficient agreement’ standard and the
17 circular reasoning needed to arrive at a firearm identification has no acceptance in
18 mainstream science.”). General acceptance should be limited to observations consistent
19 with class characteristics: “At a foundational level, beyond comparing class characteristics
20 forensic toolmark practice lacks adequate scientific underpinning and the confidence of the
21 scientific community as whole.” *Ross*, 68 Misc. 3d at 917. Thus, like the others, this factor
22 favors inadmissibility.
23
24

25 Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

Conclusion

Examiner conclusions that carry foundational validity would appear limited to observing class characteristics, and saying that bullets or casings are consistent²³ with a particular class or type of firearm, and, at most, cannot be excluded as having been fired from a particular gun. Even that is problematic in terms of a capacity to mislead or cause the jury to draw unwarranted and unsupported inferences. But that is a very different thing from testimony that bullets and casings are a “match” to a particular firearm, or were, in fact, fired from an individual firearm based on a scientific or quasi-scientific comparison analysis of toolmarks. *See Adams*, 444 F. Supp. 3d at 1267 (“It is possible that the AFTE method could be expressed in scientific terms, but I have not seen it done in this case, nor elsewhere.”). The latter should, unequivocally, be prohibited: “[I]n light of the inability of the published studies to establish an error rate, the absence of an objective standard for identification, and the lack of general acceptance of the foundational validity of the field outside of the community of practitioners within the field—reliable principles and methods do not adequately support the theory that a firearms examiner can identify a particular firearm as having fired a particular bullet or cartridge casing.” *Tibbs*, 2019 D.C. Super. LEXIS 9, at *75.

²³ That is different than opining that markings on a bullet or cartridge case are “consistent” with those on other cartridge components or with being fired from an individual firearm, which remains problematic for the reasons cited above.

1 It is particularly dangerous here, where the government will attempt to prove Mr.
2 Cloud's guilt through several other highly fallible forms of testimony that have already been
3 responsible for many wrongful convictions: eyewitness identification and cooperator. Even
4 standard jury instructions caution the factfinder to treat such evidence with great caution²⁴.
5 Coupling that proof with more evidence of dubious reliability risks erroneous, unjust results,
6 especially given the tendencies of jurors to overly credit forensic or scientific testimony, as
7 Judge Rakoff recognized in *Glynn*, 578 F. Supp. 2d at 571:

8 [O]nce expert testimony is admitted into evidence, juries are required to evaluate
9 the expert's testimony and decide what weight to accord it, but are necessarily
10 handicapped in doing so by their own lack of expertise. There is therefore a special
11 need in such circumstances for the Court, if it admits such testimony at all, to limit
12 the degree of confidence which the expert is reasonably permitted to espouse.

13 The hallmark of religion is faith, whereas the hallmarks of science include proof,
14 testability and reproducibility. When it comes to firearms identification, particularly at the
15 level of a claimed match to a particular firearm, AFTE dogma, which is akin to religion,
16 should not prevail in the courtroom. Faith – belief and supposition – can be a powerful
17 thing, but before admitting evidence to condemn a man possibly for the rest of his life,
18 courts require more. In the absence of evidence of reliability, the Court should exclude
19

20
21
22
23 ²⁴ See e.g., Ninth Circuit Model Instruction § 4.9; First Circuit Pattern Jury Instruction § 2.22
24 (eyewitness identification testimony “must be received with caution and scrutinized with care”); Sand et al.,
25 *Modern Federal Jury Instructions*, § 7-5; see also *Perry v. New Hampshire*, 565 U.S. 228, 246-47 & n.7 (2012); *United States v. Tirouda*, 394 F.3d 683, 687-88 (9th Cir. 2005).
Motion to Exclude or Alternatively Limit Toolmark Identification Evidence

1 firearms identification testimony *in toto*, or limit it consistent with the arguments and
2 authorities discussed above. Mr. Cloud respectfully asks the Court to grant the motion.
3

4 Dated: June 15, 2021

5 Respectfully Submitted,

6 S/Jeremy B. Sporn

7 Jeremy B. Sporn, NY 4779310
8 Attorneys for James D. Cloud
9 Federal Defenders of
10 Eastern Washington and Idaho
11 306 East Chestnut Avenue
12 Yakima, Washington 98901
13 (509) 248-8920
14 Email: Jeremy_Sporn@fd.org

15 **CERTIFICATE OF SERVICE**

16 I hereby certify that on June 15, 2021, I electronically filed the foregoing with the
17 Clerk of the Court using the CM/ECF System which will send notification of such filing
18 to the following: THOMAS J. HANLON and RICHARD C. BURSON, Assistant
19 United States Attorneys.
20

21 S/Jeremy B. Sporn

22 Jeremy B. Sporn, NY 4779310
23 Attorneys for James D. Cloud
24 Federal Defenders of
25 Eastern Washington and Idaho
306 East Chestnut Avenue
Yakima, Washington 98901
(509) 248-8920
Email: Jeremy_Sporn@fd.org

Motion to Exclude or Alternatively Limit Toolmark Identification Evidence